



Caldwell, D., & Welton, N. (2016). Approaches for synthesising complex mental health interventions in meta-analysis. *Evidence-Based Mental Health*, 19(1), 16-21. <https://doi.org/10.1136/eb-2015-102275>

Peer reviewed version

License (if available):
Unspecified

Link to published version (if available):
[10.1136/eb-2015-102275](https://doi.org/10.1136/eb-2015-102275)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via BMJ Publishing Group at <http://dx.doi.org/10.1136/eb-2015-102275>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/eb-terms/>

Approaches for synthesising complex mental-health interventions in meta-analysis

Deborah M. Caldwell¹ and Nicky J. Welton¹

1. School of Social and Community Medicine, University of Bristol, 39 Whatley Road, Bristol, UK

Correspondence to d.m.caldwell@bristol.ac.uk

Introduction

According to the Centre for Evidence-Based Medicine's hierarchies of evidence, systematic reviews of homogenous well-conducted randomised controlled trials (RCTs) provide the best evidence for evaluating intervention effectiveness [1]. Homogeneity implies that each study included in a meta-analysis is estimating a single, true underlying relative intervention effect, so that any differences in estimates between studies are due to sampling error alone [2]. If homogeneity does not hold, a single, fixed treatment effect meta-analysis model should not be assumed and a random effects model may be more appropriate [2] [3]. In systematic reviews of interventions for mental health, it could be argued that the homogeneity assumption is unlikely to hold in general. This is because the *clinical* variation observed across patient populations, therapist fidelity, intervention and comparator conditions, and outcomes (both patient- and clinician-reported) can give rise to *statistical* heterogeneity in meta-analysis. Indeed, heterogeneity might be considered inevitable [4].

In systematic reviews of mental health interventions, the presence of statistical heterogeneity may be attributable to the *complexity* of the intervention being evaluated leading to potentially important differences across studies. 'Complexity' itself is a contested term [5], however the MRC have described the characteristics of complex interventions as having:

- A number of interacting components within the experimental and control interventions,
- A number and difficulty of behaviours required by those delivering or receiving the intervention,
- A number of groups or organisational levels targeted by the intervention
- A number and variability of outcomes
- A degree of flexibility or tailoring of the intervention permitted [6].

These characteristics may, or may not, be present in every complex intervention. Petticrew *et al* classify characteristics of complexity as (1) those, which relate to the intervention itself (such as multiple interacting components and flexibility of implementation) and (2) those which relate to the interventions' causal pathway (such as interaction with context, multiple mediators and moderators of effect) [7].

Strategies to handle complex interventions in meta-analysis range from "lumping" all interventions together [9] to sophisticated statistical modelling techniques [34]. The aim of this paper is to give an

overview of the different analytical strategies suggested for incorporating intervention complexity in a systematic review, illustrated using a subset of studies from a Cochrane review examining psychological therapies for reducing depressive symptoms post-coronary heart disease [8]. We consider complexity only as it relates to the intervention and conceptualise a complex intervention as one, which has multiple, potentially interacting components. This is the most common interpretation [9]. Interested readers are referred to a special edition of the *Journal of Clinical Epidemiology* for consideration of strategies for handling to handle other aspects of complexity in systematic reviews [10].

Formulating the research question – lumping or splitting?

The analytical strategy for synthesising complex interventions should be pre-specified and begins with the formulation of a sensible research question, which in turn depends on the purpose of the review [11] [12]. The specification of a review's objectives shapes whether the analytical strategy will “lump” or “split” interventions. For example, ‘in principle’ research questions such as “do psychological therapies (as a whole), reduce depression after coronary heart disease?” might take a lumping approach to analysis, since this question seeks to understand effectiveness in general. However, when complex interventions are ‘lumped’ together to form a single comparator, any between interventions variation is masked and is likely to manifest as increased, but unexplained, heterogeneity. Of course, the decision to lump interventions may also be taken for practical reasons, such as when there are few eligible studies for inclusion in the review.

Consider Figure 1, which is adapted from a Cochrane review of 36 psychological interventions for coronary heart disease [8]. The outcome of interest here is reduction in depressive symptoms, for which 11 studies were included. The comparison is any psychological intervention vs control, where control is defined as standard care/ treatment as usual (TAU). A fixed-effect meta-analysis was conducted by the authors and a standardised mean reduction of -0.18 (95% Confidence Interval [CI] - 0.24 to -0.12) suggests that psychological interventions may affect a modest reduction in depression post-coronary heart disease. However, the p-value for the chi-square statistic provides extremely strong evidence against the null hypothesis of homogeneity (i.e. that interventions are estimating a single underlying treatment effect).

Figure 1 about here

Figure 2 about here

The I^2 statistic suggests that 75% of the variation between studies is attributable to heterogeneity and not chance. To account appropriately for the observed between-study variation a random-effect meta-analysis may have been more appropriate [2]. However, this would still only answer an ‘in principle’ question of general effectiveness and results would not enable a clinician to select a specific psychological intervention for their patient. The meaningful analysis of complex interventions can therefore pose problems if a ‘lumped’ approach is followed. Further exploration can be achieved by conducting an *a priori* specified sub-group analysis [3]. Figure 2 shows subgroup analysis by mode of therapy delivery, however this does not appear to explain the observed heterogeneity (Individual therapy $I^2= 77\%$; Group therapy $I^2=86\%$), and the test for subgroup

differences is non-significant ($p=0.31$). In principle, the interventions could be further sub-grouped such as “individual + weekly meetings” or “group + weekly meetings + telephone support”. However, caution should be exercised since such analyses may suffer from low power due to the small number of included studies in each grouping. If the purpose of a review is to investigate which type of psychological intervention is effective, or which intervention characteristics are effective, then a review which categorises the intervention characteristics and ‘splits’ the analysis by intervention type may be the more appropriate and robust strategy. This can either be achieved as a series of separate reviews [13-16] or as separate analyses within the same review [17].

Categorisation of intervention characteristics

There are a number of ways in which a splitting approach can be applied for meta-analyses of complex interventions. One possibility is to use the theoretical underpinning of the interventions to construct “clinically meaningful units”, which should be specified *a priori* [18]. In clinical psychology this might include classification by intervention modality such as cognitive behavioural therapy, humanistic therapy, or behavioural therapy. In reviews of mental ill-health prevention, interventions could be grouped by psychological or behavioural theory, such as the theory of planned behaviour, health belief model, social-cognitive theory, and so on [19] [20]. In Figure 3 the psychological interventions for coronary heart disease have been categorised according to intervention modality. The information was obtained from the Characteristics of Studies tables included in the original Cochrane review. The three modalities were cognitive behavioural therapy (CBT), behavioural therapy (BT) and counselling based interventions. Of these, we note only behavioural therapy was associated with a reduction in depression and the I^2 is 0%. However, the between study heterogeneity is still very high for CBT and counselling, and further investigation is warranted (note that estimates of heterogeneity become problematic when few studies are involved). One could disaggregate the intervention modalities further; for example, under CBT one might be interested in problem solving therapies or rational-emotive behavioural therapies [21]. Note however that a balance needs to be found between a sufficiently detailed categorisation that can explain heterogeneity and sufficient numbers of studies for statistical power and to avoid spurious findings.

Figure 3 about here

Components-based network meta-analysis

The obvious difficulty for standard, pairwise meta-analyses which seek to disaggregate complex interventions is that there are typically too few studies to allow clinically useful ‘splitting’ and inevitably, some degree of aggregation is needed for a meta-analysis to be conducted. In network meta-analyses, however, the potential for disaggregating the intervention is more promising [18]. A network meta-analysis (NMA) is an extension of traditional pair-wise meta-analysis to include multiple interventions, as long as these interventions form a connected network of evidence (see Figure 4, which shows a “star network” where all interventions have been compared with the same common comparator - TAU). A key advantage of NMA is that it produces summary estimates of relative effectiveness regardless of whether interventions have been compared directly and ranks them according to the outcome measured (e.g., effectiveness or safety). An additional advantage of

NMA is that it allows more studies to be combined, as long as they connect to the network (eg CBT vs Counselling studies could be added to Figure 4), bringing increased precision in the estimated intervention effects and the potential to explore statistical heterogeneity. For further details on the statistical methodology readers should see [22-24] and for a discussion of the implications for systematic review methodology see [25].

Figure 4 about here

The ‘clinically meaningful unit’ classification approach explored above has been applied to a network of psychotherapies for treating depression [26], treating acute depression in primary care [27] and psychotherapies for panic disorder [28]. Indeed, the ‘clinically meaningful unit’ analysis presented above in Figure 3 could be re-analysed as an NMA with three interventions, and sharing a common heterogeneity parameter. Figure 4 depicts the network structure for this analysis; note that all active psychological interventions are compared to the ‘usual care’ node forming a star-shaped network. Just as in the pairwise meta-analyses above, it is assumed that the standard/ TAU comparators are similar enough to be combined with the additional assumption that this must now apply across all interventions [22].

In Figure 5 the findings from the NMA are reported not only for the comparisons on which there is direct evidence but also for those where it is absent e.g BT vs CBT. There is substantial uncertainty surrounding the pairwise estimates of intervention effect, and the only comparison that reaches conventional statistical significance is BT vs TAU (note the NMA was performed in a Bayesian framework, which accounts for the wider confidence intervals when compared to Fig. 3). On the basis of this NMA, CBT is ranked 3rd (95% CIs 1st to 4th) best in terms of reducing depressive symptoms, BT is ranked 1st (95% CIs 1st to 3rd), and counselling is ranked 2nd (95% CIs 1st to 4th). Treatment as usual is the ‘worst’ intervention. Note the confidence intervals around the rankings reflect the considerable uncertainty observed in the effect estimates. In NMA a single between study heterogeneity parameter is typically assumed [22]. Here the estimate of τ^2 is 0.11 which might be considered to represent a moderate level of heterogeneity. Note that the NMA assumes that the heterogeneity is the same regardless of which comparison is being made. This may not be appropriate here, since we found more heterogeneity in the CBT vs TAU comparison than for the other comparisons (Fig. 3). This suggests that the CBT classification may be too broad to capture the complex nature of CBT interventions.

Figure 5 about here

Multi-components-based network meta-analysis

Within a NMA framework, the analyst has greater flexibility to evaluate complex multi-component interventions and to investigate whether interventions with a particular component(s) are more likely to be effective. Components are defined as the “active ingredients”, ‘intervention techniques’, or ‘elements of an intervention that have the potential to causally influence outcomes’ [9]. As such they may be classified on practical elements e.g. activities, mode of delivery, setting and/ or on theoretical underpinnings of the intervention. If there are common components across all interventions in the network, the components effectively become the intervention ‘nodes’ in the network and a NMA can be conducted. Figure 6 represents a multi-components-based network plot

for the coronary heart disease example. Welton *et al* [29] conducted a components-based NMA for the coronary heart disease network. Interventions were classified according to five key components; educational, behavioural, cognitive, relaxation, and psychosocial support. Describing their model as a meta-regression based extension to NMA, three models were evaluated in a Bayesian framework: (1) an additive main effects model which assumes that the effect of each component adds (i.e. no synergistic or antagonistic effects), (2) a 2-way interaction model (allowing pairs of components to have either a bigger or smaller effect than would be expected from the sum of their effects alone) and (3) a full interaction model for interventions described as having >2 components (e.g. cognitive+behavioral+support). To illustrate, their results for the depression outcome are shown in Figure 7 for the main effects additive model. This analysis answers the question “which intervention component has the greatest probability of being most effective?” Compared to the broader categorisation used in Fig. 5, having broken down interventions into their component parts heterogeneity is now reduced; $Tau^2 = 0.03$. There is some evidence that an intervention with a cognitive and/or behavioural component(s) was associated with a reduction in standardized mean depression score; for the cognitive component the pooled SMD was -0.26 (95% credible interval: -0.55 to 0.02) and for behavioural it was SMD -0.24 (95% credible interval: -0.42 to -0.06).

Figure 6 about here

Figure 7 about here

Discussion

Component-based network meta-analysis is an option for the synthesis of complex interventions in the presence of heterogeneity. Of course, intervention categorisation is only one dimension contributing to heterogeneity in meta-analyses of complex interventions. In the above example, heterogeneity was explained by intervention definition but this may not be the case for all examples, where additional factors may cause residual heterogeneity (for example an imbalance of effect modifiers across studies). A possible source of confounding here is the control intervention. In RCTs in clinical psychology and psychiatry control interventions may take several forms – waiting list controls are common as are no intervention controls [30]. A psychological placebo, where the intervention is regarded as inactive by the researchers but is judged as active by the participants, may be used. Similarly an attention placebo could be used where the control mimics the theoretically inactive elements of an intervention, but not the active elements [31, 32]. Reviewers of complex interventions should also be mindful that treatment as usual and standard care may differ across settings, contexts and countries, even though systematic reviews have traditionally lumped these into a single control [33, 34]. Unfortunately, due to the small number of studies in the psychotherapy for coronary heart disease meta-analysis, further disaggregating by control intervention is of questionable value.

Component based systematic reviews are becoming increasingly common as analysts realise the importance of identifying and investigating heterogeneity, regardless of its inevitability [35, 36]. However, one difficulty in a components based approach is the identification of distinct components from the published literature [37]. Complex interventions may not be described in sufficient detail to allow dismantling of key ingredients. Recent reporting guideline initiatives, such as CReDICI 2

(Criteria for Reporting the Development and Evaluation of Complex Interventions in healthcare: revised guideline) [38] and CONSORT-SPI (Social and Psychological Interventions) [39] seek to address this. The MRCs recent guidance on process evaluations may also help identification of components, and assessment of delivery and fidelity of complex interventions [40]. How to classify complex interventions and disaggregate the multiple interacting components within them is an area of ongoing interest. Several taxonomies have been developed; some designed for use in specific clinical areas and others are generic [41]. Further research is needed to assess the application of taxonomies across clinical areas. Logic models describing the mechanisms of action and casual pathways of interventions are increasingly used to structure systematic reviews of complex interventions [42] and could also be used to inform the classification of intervention components. What is clear, however, is that whichever approach the analyst chooses to categorise interventions it is desirable that components be specified *a priori*, and published in a protocol before data extraction to avoid data-driven decisions and reduce the likelihood of spurious findings.

Funding acknowledgements

DMC is supported by a Medical Research Council Population Health Scientist fellowship award G0902118.

This work was undertaken with the support of The Centre for the Development and Evaluation of Complex Interventions for Public Health Improvement (DECIPHer), a UKCRC Public Health Research Centre of Excellence. Joint funding (MR/K0232331/1) from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, the Welsh Government and the Wellcome Trust, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged.

References

1. Centre for Evidence BAsed Medicine *Levels of Evidence*. 2015 [cited 2015 15-10-2015]; Available from: <http://www.cebm.net>.
2. Nikolakopoulou, A., D. Mavridis, and G. Salanti, *Demystifying fixed and random effects meta-analysis*. Evid Based Ment Health, 2014. **17**(2): p. 53-7.
3. Higgins, J., S. Green, and editors, *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. 2011, The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
4. Higgins, J.P.T., S.G. Thompson, and D.J. Spiegelhalter, *A re-evaluation of random-effects meta-analysis*. Journal of the Royal Statistical Society. Series A, (Statistics in Society), 2009. **172**(1): p. 137-159.
5. Wong, G., *Is complexity just too complex?* J Clin Epidemiol, 2013. **66**(11): p. 1199-201.
6. Craig, P., et al., *Developing and evaluating complex interventions: the new Medical Research Council guidance*. BMJ, 2008. **337**.
7. Petticrew, M., et al., *Complex interventions and their implications for systematic reviews: a pragmatic approach*. J Clin Epidemiol, 2013. **66**(11): p. 1209-14.
8. Rees, K., et al., *Psychological interventions for coronary heart disease*. Cochrane Database Syst Rev, 2004(2): p. Cd002902.
9. Kühne, F., et al., *Conceptual decomposition of complex health care interventions for evidence synthesis: a literature review*. Journal of Evaluation in Clinical Practice, 2015. **21**(5): p. 817-823.

10. Tugwell, P., J.A. Knottnerus, and L. Idzerda, *Complex interventions—how should systematic reviews of their impact differ from reviews of simple or complicated interventions?* Journal of Clinical Epidemiology, 2013. **66**(11): p. 1195-1196.
11. Squires, J.E., J.C. Valentine, and J.M. Grimshaw, *Systematic reviews of complex interventions: framing the review question.* Journal of Clinical Epidemiology, 2013. **66**(11): p. 1215-1222.
12. Anderson, L.M., et al., *Introducing a series of methodological articles on considering complexity in systematic reviews of interventions.* Journal of Clinical Epidemiology, 2013. **66**(11): p. 1205-1208.
13. Churchill R, M.T., Furukawa TA, Caldwell DM, Davies P, Jones H, Shinohara K, Imai H, Lewis G, Hunot V. , *'Third wave' cognitive and behavioural therapies versus treatment as usual for depression.* . Cochrane Database of Systematic Reviews 2013, , 2013. **Issue 10.** : p. Art. No.: CD008705. DOI: 10.1002/14651858.CD008705.pub2.
14. Hunot V, M.T., Caldwell DM, Furukawa TA, Davies P, Jones H, Honyashiki M, Chen P, Lewis G, Churchill R., *'Third wave' cognitive and behavioural therapies versus other psychological therapies for depression.* Cochrane Database of Systematic Reviews 2013,, 2013. **Issue 10. Art. No.: CD008704.** : p. DOI: 10.1002/14651858.CD008704.pub2.
15. Shinohara K, H.M., Imai H, Hunot V, Caldwell DM, Davies P, Moore THM, Furukawa TA, Churchill R. , *Behavioural therapies versus other psychological therapies for depression.* . Cochrane Database of Systematic Reviews 2013,, 2013. **Issue 10.:** p. Art. No.: CD008696. DOI: 10.1002/14651858.CD008696.pub2.
16. Caldwell Deborah, H.V., Moore Theresa HM, Davies Philippa, Jones Hannah, Lewis Glyn, Churchill Rachel, *Behavioural therapies versus treatment as usual for depression.* Cochrane Database of Systematic Reviews, 2010. **2010**(9): p. DOI: 10.1002/14651858.CD008697.
17. Petticrew, M., et al., *Synthesizing evidence on complex interventions: how meta-analytical, qualitative, and mixed-method approaches can contribute.* Journal of Clinical Epidemiology, 2013. **66**(11): p. 1230-1243.
18. Melendez-Torres, G.J., C. Bonell, and J. Thomas, *Emergent approaches to the meta-analysis of multiple heterogeneous complex interventions.* BMC Med Res Methodol, 2015. **15**: p. 47.
19. Faggiano, F., et al., *School-based prevention for illicit drugs use: a systematic review.* Prev Med, 2008. **46**(5): p. 385-96.
20. Michie, S. and A. Prestwich, *Are interventions theory-based? Development of a theory coding scheme.* Health Psychol, 2010. **29**(1): p. 1-8.
21. Churchill, R., et al., *Cognitive behavioural therapies versus other psychological therapies for depression.* The Cochrane database of systematic reviews, 2010(9): p. CD008698.
22. Mavridis, D., et al., *A primer on network meta-analysis with emphasis on mental health.* Evidence Based Mental Health, 2015. **18**(2): p. 40-46.
23. Dias S, A.A., Sutton AJ, Welton NJ, *Evidence Synthesis for Decision Making 2: A Generalized Linear Modeling Framework for Pairwise and Network Meta-analysis of Randomized Controlled Trials.* Medical Decision Making, 2013. **33**: p. 607-617.
24. Dias, S., et al., *Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials.* Medical Decision Making, 2013. **33**(5): p. 641-56.
25. Caldwell, D.M., *An overview of conducting systematic reviews with network meta-analysis.* Syst Rev, 2014. **3**: p. 109.
26. Barth J, M.T., Gerger H, Nuesch E, Trelle S, Znoj H, Juni P, Cuijpers P., *Comparative Efficacy of Seven Psychotherapeutic Interventions for Patients with Depression: A Network Meta-Analysis.* PLOS Medicine, 2013.
27. Linde, K., et al., *Effectiveness of Psychological Treatments for Depressive Disorders in Primary Care: Systematic Review and Meta-Analysis.* The Annals of Family Medicine, 2015. **13**(1): p. 56-68.
28. Pompili A, F.T., Imai H, Tajika A, Efthimiou O, Salanti G, *Psychological therapies for panic disorder with or without agoraphobia in adults.* Cochrane Database of Systematic Reviews, In press.
29. Welton, N.J., et al., *Mixed treatment comparison meta-analysis of complex interventions: psychological interventions in coronary heart disease.* Am J Epidemiol, 2009. **169**(9): p. 1158-65.
30. Button, K.S. and M.R. Munafò, *Addressing risk of bias in trials of cognitive behavioral therapy.* Shanghai Archives of Psychiatry, 2015. **27**(3): p. 144-148.
31. Furukawa, T.A., et al., *Waiting list may be a nocebo condition in psychotherapy trials: a contribution from network meta-analysis.* Acta Psychiatrica Scandinavica, 2014. **130**(3): p. 181-192.

32. Zhu, Z., et al., *Comparison of psychological placebo and waiting list control conditions in the assessment of cognitive behavioral therapy for the treatment of generalized anxiety disorder: a meta-analysis*. Shanghai Archives of Psychiatry, 2014. 26(6): p. 319-331.
33. Cuijpers, P., et al., *Efficacy of cognitive-behavioural therapy and other psychological treatments for adult depression: meta-analytic study of publication bias*. Br J Psychiatry, 2010. 196: p. 173-8.
34. Jakobsen, J.C., et al., *The effects of cognitive therapy versus 'no intervention' for major depressive disorder*. PloS one, 2011. 6(12): p. e28299.
35. Greaves, C.J., et al., *Systematic review of reviews of intervention components associated with increased effectiveness in dietary & physical activity interventions*. BMC Public Health, 2011. 11: 119.
36. Squires, J.E., et al., *Are multifaceted interventions more effective than single-component interventions in changing health-care professionals' behaviours? An overview of systematic reviews*. Implement Sci, 2014. 9: p. 152.
37. Guise, J.-M., et al., *Agency for Healthcare Research and Quality Evidence-based Practice Center methods for systematically reviewing complex multicomponent health care interventions*. Journal of Clinical Epidemiology, 2014. 67(11): p. 1181-1191.
38. Mohler, R., S. Kopke, and G. Meyer, *Criteria for Reporting the Development and Evaluation of Complex Interventions in healthcare: revised guideline (CRDeCI 2)*. Trials, 2015. 16: p. 204.
39. Gardner, F., et al., *Editorial Perspective: The need for new guidelines to improve the reporting of trials in child and adolescent mental health*. Journal of Child Psychology and Psychiatry, 2013. 54(7): 810-12.
40. Moore, G.F., et al., *Process evaluation of complex interventions: Medical Research Council guidance*. BMJ, 2015. 350.
41. Michie, S., et al., *The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions*. Ann Behav Med, 2013. 46(1): p. 81-95.
42. Anderson, L.M., et al., *Using logic models to capture complexity in systematic reviews*. Research Synthesis Methods, 2011. 2(1): p. 33-42.
43. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.
44. Available from <http://www.openbugs.net/>